Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.25, NO.2, Apr. 2015 ISSN 2288-2715(Online)
http://dx.doi.org/10.13089/JKIISC.2015.25.2.293

스미싱 범죄 프로파일링 모델 설계

정 영 호,[1†] 이 국 헌,[2] 이 상 진[2‡]
[1]경찰대학 국제사이버범죄연구센터, [2]고려대학교 디지털포렌식 연구센터

Designing SMS Phishing Profiling Model

Youngho Jeong,[1†] Kukheon Lee,[2] Sangjin Lee[2‡]
[1]International Cybercrime Research Center, Korean National Police University,
[2]Digital Forensic Research Center, Korea University

요 약

스미싱 범죄 피해 사례에서 수집할 수 있는 공격정보들을 이용하여, 범죄수사에 사용하는 프로파일링 기법을 응용한 스미싱 범죄 프로파일링 모델을 제안한다. 기존에 수사기관에서는 apk 파일의 해시를 이용한 시그니처 분석과 코드 내 삽입된 C&C IP 분석방법을 사용하였으나, 시그니처의 다변화와 코드 난독화로 인해 그 활용도가 낮아졌다. 실제 수사기관에 접수된 169건의 피해사례의 분석을 통해, apk 파일 내 인증서 파일 일련번호의 재사용이 151건(89%), 퍼미션 파일의 재사용은 136건(80%)에 달한다는 점에 착안, 인증서 파일의 일련번호와 퍼미션 파일의 해시를 중심으로 한 스미싱 프로파일링 모델을 설계하여 범죄를 군집화하여 기존의 해시 기반 군집화 방법을 보완하였고, 코드 유사도 검증을 통하여 추가로 신뢰성을 확보하였다.

ABSTRACT

With the attack information collected during SMS phishing investigation, this paper will propose SMS phishing profiling model applying criminal profiling. Law enforcement agencies have used signature analysis by apk file hash and analysis of C&C IP address inserted in the malware. However, recently law enforcement agencies are facing the challenges such as signature diversification or code obfuscation. In order to overcome these problems, this paper examined 169 criminal cases and found out that 89% of serial number in cert.rsa and 80% of permission file was reused in different cases. Therefore, the proposed SMS phishing profiling model is mainly based on signature serial number and permission file hash. In addition, this model complements the conventional file hash clustering method and uses code similarity verification to ensure reliability.

Keywords: Cyber-crime, Profiling, Smishing

I. 서 론

경찰청이 발표한 통계에 의하면, 스미싱 범죄는 2012년에 피해 접수 건수 2,182건, 피해금액 5억 6,900만원에서 2013년에는 피해 접수 건수 신고는 29,761건으로 피해액 57억원에 이를 정도로 급증하였다[1]. 이에 2013년 12월, 금융위원회 주관 범부처 대책협의회에서는 '신·변종 전기통신금융사기 피해방지 종합대책'을 발표하였다[2]. 하지만 2014년 1월부터 5월까지 스미싱 피해 접수 또한 3,097건(경찰청), 스미싱 탐지(차단) 건수는 총 97만 4302건(KISA)으로, 그 시도는 전혀 줄어들지 않고 있다.

일반 강력범죄의 경우 한 건의 범죄에선 범인(범죄조직)과 피해자가 같은 범죄장소에 위치하지만, 스미싱과 같은 사이버범죄는 범죄를 저지르는 장소에

접수일(2014년 11월 26일), 게재확정일(2015년 1월 22일)
† 주저자, knp05@police.go.kr
‡ 교신저자, sangjin@korea.ac.kr(Corresponding author)

관계없이 피해자들이 전국에 퍼져 있다는 문제가 있다. 또한 그 수사가 전국의 경찰관서에서 개별적으로 진행되어 행정력이 낭비되고 있다.

이러한 문제를 해결하기 위해 피해자가 신고하면서 제출한 apk 파일의 해시 값과 코드 내에 나타나 있는 IP 주소를 통하여 범죄 행위를 클러스터링(clustering)하였다. 그러나 난독화 기법이 사용된 이후 IP 주소를 빠르게 찾기 어려워지면서 클러스터링하는 것이 어려워졌다.

본 논문은 지금까지 제시된 스미싱 범죄의 클러스터링에 대한 문제점을 분석하여, 사이버범죄 프로파일링 기법을 이용한 스미싱 범죄 프로파일링 모델을 제안한다. 본 논문의 2장과 3장에서는 스미싱 범죄 현황과 기존 사이버범죄 프로파일링, 모바일 악성코드의 분류방법에 대한 선행 연구 내용을 살펴보며, 4장에서는 apk 파일의 인증서 일련번호, 퍼미션 파일의 해시 값 등의 정보를 기반으로 한 스미싱 프로파일링 모델을 설계하고, 사례 분석을 통해 효과성을 검증한다.

II. 배경 지식

2.1 스미싱의 정의

스미싱(smishing)은 인터넷 보안기업 맥아피(McAfee)가 2006년에 처음 사용한 용어이다. 문자메시지(SMS)와 피싱(phishing)의 합성어이며, 휴대전화 문자메시지 서비스를 통해 발송되는 피싱 공격을 의미한다[3]. 공격자는 Web 접속이 가능한 피해자의 휴대전화에 URL 주소가 포함된 문자메시지를 발송하고, 피해자가 그 링크를 클릭하면 트로이목마가 설치되는 원리이다. 경찰청 사이버안전국에서는 '무료쿠폰 제공', '돌잔치 초대장', '모바일청첩장' 등을 내용으로 하는 문자메시지 내 인터넷주소를 클릭하면 악성코드가 스마트폰에 설치되어, 피해자가 모르는 사이에 소액결제 피해나 개인·금융정보를 탈취당하는 사이버범죄 행위를 스미싱으로 정의하고 있다[4].

2.2 스미싱 발생 유형

국내에서는 2007년 초 국내 금융기관의 인터넷뱅킹 사이트를 모방한 '피싱사이트' URL을 발송하는 수법이 처음 발견되었다. 피싱사이트로 연결되는 URL과 함께 '보안승급강화' 등의 메시지를 발송하고, 이에 속은 피해자가 해당 사이트에 접속하여 금융정보를 입력하도록 하여 금원 및 개인정보를 탈취하는 수법이었다.

2012년에 발생한 스미싱 유형은 안드로이드 스마트폰 운영체제의 '알 수 없는 소스' 옵션을 해제하면 나타나는 취약성을 이용한 공격 기법이다. 범죄조직은 '무료쿠폰'이나 '이번달 보험료 미환급', '안드로이드 업데이트', '주민번호 이용내역 확인', '법원등기확인', '모바일 청첩장', '돌잔치 초대' 등 사회공학 기법을 활용한 문자메시지를 발송, 피해자들이 단축 URL 주소가 포함된 링크를 클릭하게 함으로써 피해자들의 스마트폰에 악성 애플리케이션(apk) 파일을 설치하였다. 이를 통해 피해자에게 발송되는 문자메시지를 가로채고, 미리 준비해둔 유출 개인정보를 활용하여 소액결제를 함으로써 범행을 하였다.

본 연구에서는 이와 같은 사회공학 기법을 이용한 스미싱 범죄를 프로파일링 모델 설계의 대상으로 삼는다. 또한 최근에는 악성 애플리케이션을 이용, 개인의 통화기록과 위치정보, 사진 등의 모든 자료를 탈취하는 '스파이앱', 정상적인 스마트폰뱅킹 애플리케이션을 위·변조하여 금융정보를 요구하는 등 다양한 형태의 스미싱이 나타나고 있다.

III. 선행연구

3.1 모바일 악성코드의 분류 모델

스미싱 범죄가 급증하면서, 모바일 악성코드의 분류 방법과 범죄에 이용된 스미싱의 예방·차단 방법에 대한 연구가 활발해졌다.

윤재성 등[5]은 모바일 악성코드의 행위를 기반으로 한 Andro-profiler 방식을 제안하였고, 박창욱 등[6]은 퍼지해시를 이용한 유사 악성코드 분류모델 방식을 제안하였으며. 박재우 등[7]은 문자열과 API를 이용한 악성코드 자동 분류 시스템을 제안하였다. 해외에서는 퍼미션 기반의 모바일 악성코드 탐지를 위한 방식이 주로 연구되었다[8].

또한 주춘경 등[9]은 개인식별화된 SMS 발송을 통한 스팸식별 방식을 제안하여 스미싱을 예방하는 방법을, 장상근[10]은 모바일 악성코드 사례를 분석하여 악성코드 진단 방법을 제안하였다.

위 연구들은 악성코드의 시그니처, 문자열, API 등의 파일 내부 정보를 통한 악성코드의 분류방법을 제안한 논문들이다. 그러나 범죄조직이 코드 암호화

및 난독화 기법을 사용하면서, 분석 시간이 장기화되고 수사가 지연되고 있다.

본 논문에서는 파일 내부 정보를 이용한 분류모델은 간략화하고, 범죄조직이 이용한 공격정보(SMS, 단축URL)와 피해자정보를 결합하여, 악성코드 기반 분류모델로는 분류할 수 없는 범죄 피해사례별 군집화 및 프로파일링 기법을 연구한다.

3.2 프로파일링의 정의와 개념

전통적인 범죄자 프로파일링은 범행 현장에서 범죄자가 나타낸 범행 수법이나 여러 행동들의 분석을 통해서 범죄자의 사회인구학적인 배경 특성(나이, 교육 수준, 직업, 주거지, 결혼 여부, 성격 특성, 심리적/정신적 장애, 범죄 경력, 피해자와의 관계 등)을 추론하여 범인 검거에 기여하는 수사기법을 일컫는다[11]. 또 다른 개념으로는, 프로파일링이란 범죄현장에서 발견되는 여러 가지 형태의 증거를 법과학적, 심리학적 방법을 통하여 분석하고 이를 통하여 해당 증거를 남긴 범죄자의 성격 및 행동유형을 확인하는 수사방법으로, 인간의 행동유형(Behavior)은 시간과 환경에 약간의 영향을 받지만 개인의 성격을 반영하는 근본적인 행동을 변화하지 않으며 범죄 행동 또한 인간의 행동이기 때문에 행위자의 성격과 이를 반영한 일정한 행동패턴을 갖는 이론적 전제에서 출발하며[12] 이러한 프로파일링은 과학수사의 한 분야로 최근 인정을 받고 있다.

3.3 사이버범죄 프로파일링

경찰청에서는 전통적 범죄 프로파일링 개념을 사이버범죄로 확장하는 사이버범죄 프로파일링을 “개별적으로 발생하는 사이버범죄 사례에서 발견되는 행위 및 수집되는 증거의 분석 등을 통하여 획득한 자료를 집중화하여 처리함으로써, 범죄예방, 수사 등의 정보를 생산하는 체계 및 해당 체계와 유기적으로 운영되는 조직 및 정책 등의 일체”로 판단한다[13].

사이버테러 프로파일링 시스템 설계방안 연구에서는 사이버범죄 프로파일링을 위한 데이터 분석과 용의자 프로파일링을 위한 데이터 분석 등을 언급하였다[12]. 사이버테러형 범죄는 전통적인 오프라인 상의 범죄와 성격이 다르고, 짧은 시간 안에 많은 횟수의 사이버공격을 통해 다수의 피해자를 만들 수 있는 점에 강조하여 모델을 설계하였다. 사이버테러형 범죄 프로파일링에서는 공격자 정보(국적/주소/고유번호 등), 공격도구 정보(이메일/도메인/C&C 등), 범죄정보(동기/목적/정보획득방법 등), 피해자정보(개인과 단체로 나누어 - 국적/연령/학력/직업, 기업의 주업무/범죄방어능력/피해인지수준 등), 사건관계인 정보(국적/연령/학력/직업 등)를 구성요소로 분류하였다.

양홍석[14]은 사이버테러 공격 중 DDoS 범죄의 프로파일링 모델 기법을 제안하였다. DDoS공격이 기존의 범죄보다 기술적인 범죄수법을 이용한 점에 주목하여, 기존의 HPP 프로젝트 항목 중 기술적 정보 항목을 확장하여 통계적 기술정보, 세부적 기술정보, artifacts 정보, 디지털지리정보에 주목하였다.

IV. 모델 설계 방안

4.1 스미싱 애플리케이션의 특징과 분석

안드로이드용 애플리케이션 파일(apk)은 윈도우 실행파일(PE)의 구조와는 다르게, 여러 파일들이 압축되어 존재하는 방식이다. 안드로이드 애플리케이션의 압축을 해제하면, 퍼미션 파일과(Androidmanifest.xml), 실행 파일(Classes. dex), resources.arsc, res 폴더, META-INF 폴더 등이 나타난다.

스미싱에 사용되는 악성 애플리케이션의 경우, 정상적인 애플리케이션에 비해 좀 더 많은 권한을 요구하고, 피해자에게 발송된 SMS를 가로채 C&C 서버로 전송하는 코드가 발견되는 것이 특징이다[9].

이렇게 스미싱 애플리케이션에서 공통적으로 발견되는 퍼미션이나 코드 내용을 기반으로 이상 행위를 탐지하고 스미싱 애플리케이션으로 분류하는 것이 이상적이지만, 이는 탐지와 분석을 위한 기법이므로 신속성과 간편성을 필요로 하는 수사를 위해서는 적절하지 않다.

사용자가 이미 금전적인 피해를 입었고, SMS와 악성코드가 이미 스마트폰에 저장되어 있는 상황에서 이 악성코드가 ‘스미싱에 해당한다’라는 것을 코드나 행위 기반으로 재검사하는 것은 불필요한 과정이다. 피해자가 제출하는 정보들 중 필요한 부분을 수집하여 빠르게 ‘기존에 벌어진 행위와 유사하거나 동일한 것인지’ 여부를 판단하는 것이 좀더 신속하게 판단할 수 있다.

이를 검증하기 위하여 2013년 2월부터 2013년 11월까지 실제로 수사기관에 접수된 피해사례 169

건을 분석하였다. 기존에 수사기관에서 군집화를 위해 사용한 방법은 apk 파일의 해시와 C&C IP 주소 분석으로, 해당 방법을 이용할 경우 169건의 피해사례는 총 152 종류 (동일 사건 16건, 약 9%)로 나타났다. 이러한 분류 방식은 집중수사나 범죄조직의 구분을 위한 군집화의 방법으로는 거의 유의미하지 못한 방법이다.

반면, 스미싱용 악성 애플리케이션 apk 파일 내에 들어있는 퍼미션 파일과 META-INF 폴더에 들어 있는 인증서 파일(CERT.rsa)를 분석한 결과 중복되어 나타나는 경우가 굉장히 많았다.

퍼미션 파일의 경우, 파일의 해시 분석만으로 전체 169건 중 2건 이상 중복되어 발견된 경우가 총 136건, 약 80%에 달했다. 또한 인증서 파일의 경우, 파일 내부에 있는 일련번호(serial number)를 분석한 결과 전체 169건 중 중복되어 발견된 경우가 총 152건, 약 89%로 높게 나타났다. 이러한 분석 결과와 스미싱 피해자들이 제출한 정보들을 토대로, 스미싱 프로파일링 모델을 설계하였다.

4.2 항목 추출 기준과 고려할 요소

스미싱 범죄의 프로파일링 모델을 설계하면서 고려해야 하는 것은 '피해자로부터 어떤 정보를 획득할 수 있는지'와 '이 정보 중 어떤 항목을 사용해야 효과적으로 사건을 분류할 수 있는지'의 문제들이다.

일단, 피해자가 휴대전화를 가지고 수사기관에 신고하면서 제출할 수 있는 정보는 휴대전화로 발송된 문자메시지 화면, 문자메시지를 클릭하여 다운로드한 악성 애플리케이션 설치 파일(apk), 범죄자로 인해 부정 결제된 소액결제 이용내역서 등 3가지이다.

이를 통해 알 수 있는 정보는 발신 전화번호, 문자 메시지 내용, 범행에 이용된 단축 URL 주소, 범행에 이용한 소액결체 업체와 실제 이용된 컨텐츠 사이트와 스마트폰에 저장된 악성 애플리케이션 파일(apk), apk 파일의 분석을 통한 파일명과 시그니처(해시 값), C&C IP 주소와 퍼미션 파일과 해시, 인증서 파일의 해시와 일련번호 등이다.

위 정보들은 범죄 프로파일링에서 사용하는 개념 중 '공격도구 정보'에 해당한다. 사이버범죄의 경우 발생 당시 공격자의 정보(국적/주소/성별)나 범죄정보(동기)는 파악하기 어렵다. 또한 특정한 집단을 대상으로 한 범죄라고 보기 어렵기 때문에, 피해자정보 또한 모델을 설계하는 데 있어 유의미한 요소로 볼 수 없다. 결국 스미싱 범죄 프로파일링은 이와 같은 공격도구 정보(기술정보)에 의존할 수밖에 없다. 다행히, 피해자가 사건을 접수할 때 위와 같은 정보들을 모두 파악할 수 있고, 수사관이 수사를 진행하며 공격자의 정보 및 사건관계인 정보를 추가로 확보할 수 있으므로 모델이 더욱 확장될 수 있다.

이러한 공격도구 정보들 중에서, 이번 연구에서 스미싱 프로파일링 모델에 이용하는 정보는 다음과 같은 총 6개의 정보이다. 문자메시지 내용, 단축 URL 주소, apk 파일의 해시, C&C 서버의 IP 주소, 퍼미션 파일의 해시, 인증서 파일의 일련번호이다.

일단, 발신 전화번호의 경우, 실제로 범인들이 본인을 추적할 수 있는 실제 전화번호를 남기는 경우가 전무하다. 또한 최근에는 악성 애플리케이션에 감염된 사용자의 휴대전화에서 제3의 피해자에게 문자메시지가 발송되는 경우도 많기 때문에, 범인의 추적을 위한 프로파일링 요소(공격자정보)로 넣기에는 적당하지 않다. 또한 범행에 이용된 소액결제 업체와 컨텐츠 사이트의 계정도 도용되거나 해킹된 경우가 많아, 전화번호와 마찬가지로 프로파일링 설계에는 적당치 않다.

Table 1. Elements of Smishing profiling

Elements	Detail
Attack tool information (technique information)	SMS Contents
	Shorten URL address
	Hash value of apk file
	Address of C&C
	Hash value of permission file
	Serial number in CERT.rsa

4.2.1 문자메시지(SMS) 내용

스미싱의 핵심은 범죄자가 사회공학 기법을 사용하여 피해자에게 문자메시지를 발송한다는 데 있다. 전술하였듯이, 스미싱 문자메시지의 흐름은 '무료쿠폰' → '돌잔치/결혼식 초대장' → '법원(검찰)출석요구서' 등으로 변화되어 왔다. 물론 범죄자들이 동일한 apk 파일을 이용하여 다른 내용의 문자메시지를 발송할 가능성은 항상 존재한다,

하지만 문자메시지 내용을 바꾸기 전까지 특정 기간 동안 전국의 피해사례를 수집할 때에는 문자메시

지 내용이 프로파일링에서 중요한 부분이다. 문자메시지의 내용이 완전히 동일하지는 않더라도, 내용을 카테고리화(예 : 무료쿠폰/청첩장/출석요구서)를 해서 전국의 피해 사례의 유사도를 판단하는 것은 중요한 작업이다. Table 2는 스미싱에 사용된 문자메시지를 7개의 카테고리로 분류한 예시이다.

Table 2. Categorizing of SMS contents used in smishing

	Contents
1	free coupon
2	invitation
3	a summons
4	personal information
5	android
6	parcel services
7	payment

4.2.2 단축 URL

스미싱용 애플리케이션을 배포하기 위해 범죄조직은 문자메시지에 악성 애플리케이션 다운로드 URL 주소를 첨부해야 한다. 하지만 URL이 긴 경우 문자메시지에 포함되지 못하는데 이를 회피하는 방법이 단축 URL(URL Shortener)의 사용이다[15].

단축 URL이 프로파일링 모델 설계에서 중요한 이유는, 단축 URL을 통하여 악성 애플리케이션이 배포되는 원래 주소를 파악할 수 있고, C&C 서버까지 추적해낼 수 있는 가능성이 있다. 또한 이 정보를 통하여 범죄조직이 이용한 IP 주소를 추적할 수 있는 가능성도 있다.

접수된 169건의 사례에서는, 피해자들이 스미싱 발신 문자를 삭제한 경우가 많아 (97건 삭제) 유의미한 분류정보로 사용할 수는 없었다. 하지만 72건의 사례를 분석한 결과, 범죄조직이 사용한 단축 URL 서비스 중 상위 5개는 Table 3과 같다.

Table 3. Shorten URL Services

Sites	Frequency of usage	Remarks
goo.gl	8	google
derpy.me	7	-
moa.so	6	-
vo.to	5	-
2url.kr	4	domestic

4.2.3 apk 파일명과 해시 값

안드로이드 애플리케이션은 개발 환경 특성상 각 애플리케이션 개발자마다 고유한 식별 정보를 가지고 있다. 또한 스미싱용 악성애플리케이션의 경우 애플리케이션을 배포하기 위한 다운로드 서버와 데이터를 수집하기 위한 악성 애플리케이션 데이터 수집 서버(C&C 서버)에 대한 주소가 코드 내에 존재한다.

실제로 수사기관에서는 집중수사를 위해 apk 파일의 해시 값(MD5)과 파일 내 코드를 분석해서 찾아낸 C&C IP 주소를 이용하였다. 그러나 해시를 통한 시그니처 분석 기법의 단점은 파일에 조금만 수정이 일어나도 그 값이 모두 변한다는 점이다. 예컨대, 루트 영역에 숨어있는 apk 파일을 꺼내거나 삭제된 apk 파일을 복구하는 경우 그 값이 완전히 달라지고, 범죄조직이 퍼미션 파일을 조금이라도 수정하거나, 인증서 파일을 변경하여 수정이 일어나는 경우 또한 그 값이 달라진다.

또한 C&C IP의 경우에도 스미싱이 발견된 초기(2012년)에는 특정 클래스에 C&C IP가 하드코딩되어 있었고 코드도 굉장히 간단하였지만, 점차 코드를 암호화 하는 경향이 나타났고, Apk-protect 등의 난독화 도구가 이용되면서 C&C 서버의 IP를 곧바로 찾아내는 일은 굉장히 어렵게 되었다. 하지만 apk 파일의 해시가 동일한 경우 동일한 범죄조직에 의한 피해사례로 볼 수 있고, apk 파일이름 또한 그 유사도의 개연성이 어느 정도 있기 때문에, 스미싱 범죄 프로파일링 모델에서 사용하는 요소로 지정하였다.

4.2.4 명령제어(C&C) 서버 IP

169건의 사례를 분석한 결과, C&C 서버의 정보가 나타난 피해 사례가 130건이다. 그러나 이 사례들이 2013년 2월부터 2013년 11월까지의 피해 사례이며, 전술하였듯이 최근에는 암호화와 난독화 등으로 인하여 주소를 곧바로 파악하기 어렵고, C&C 서버로 이메일 방식을 채택하거나, 도메인을 주기적으로 변경하는 등의 수법이 나타나고 있으므로, 이 또한 프로파일링을 위한 요소로 강하게 사용하기는 어렵다. 다만 apk 파일과 마찬가지로, C&C IP 주소를 파악할 수 있고, 그 주소가 동일하다면 역시 동일한 범죄조직의 소행으로 판단할 가능성이 높아진다. 그러므로 이 역시 프로파일링 모델의 요소로 지정한다. 일단 C&C의 형태가 IP 주소인지, 도메인

Table 4. Examples of C&C Server

	Address of C&C	Form
1	103.17.117.O	IP
2	xqqq.OOOO.net:9998	Domain
3	xingigogsi23@ OOOOO.com	E-mail

주소인지, 이메일 주소인지 분류하고, IP 주소인 경우 C 클래스까지와 도메인 서버의 국가를 그 판단 요소로 한다.

4.2.5 퍼미션 파일의 해시 값

안드로이드 애플리케이션이 특정 데이터에 접근하거나 기능을 실행하기 위해서는 각 기능에 맞는 퍼미션을 선언해야 한다. 퍼미션은 apk 파일의 압축을 해제하면 Androidmanifest.xml 파일을 통해 확인할 수 있다.

169건의 사례를 분석한 결과, apk 파일의 해시 값이 중복되어 나타난 경우는 16건에 불과하였지만, 퍼미션이 중복되어 나타난 경우는 136건으로 매우 높게 나타났다.(중복 사용 총 32종류) 퍼미션 파일이 애플리케이션 개발 단계에서 선언하고 사용된다는 점에서, 개발자들은 퍼미션 파일을 대부분 재사용한다는 가설을 세울 수 있다. 또한 중복 사용이 발견되지 않은 퍼미션 파일의 경우에도 피해사례 분석을 전국으로 확대한다면 충분히 중복된 사례를 발견할 수 있을 것이다.

Table 5의 사례는, 해시 값이 다른 5개의 apk 파일(danal.apk와 mrk.apk)내에 있는 퍼미션 파일이 동일하게 나타난 사례이다.

Table 5. Hash value(Smishing apk files and permission files)

File name	Hash value (apk)	Hash value (Androidmanifest)
danal.apk	f98807fd35fc85622d4a834414a63b9b	f6021c5cda569cce16722d19416f886f
danal.apk	ca0f3ac5ed8376cc67bd4a4671f69532	f6021c5cda569cce16722d19416f886f
danal.apk	44d6ca62e9f14db4d59b36d12127c584	f6021c5cda569cce16722d19416f886f
mRK.apk	910b4014bd50326230b36077d429d54a	f6021c5cda569cce16722d19416f886f
mrk.apk	cc2dea03108c7299fd49f41878180f41	f6021c5cda569cce16722d19416f886f

4.2.6 인증서 파일의 일련번호(Serial Number)

전술한 퍼미션 파일과 마찬가지로, 안드로이드 애플리케이션이 실행되기 위해서는 개발자의 서명이 들어가야 하고, 이 정보가 남아있는 것이 META-INF 폴더에 들어있는 CERT.RSA 파일이다. 애플리케이션을 배포하기 위한 릴리즈 파일을 생성하기 위해 사용하며, 구글 마켓에 등록하기 위해서는 필수적으로 요구된다[16]. 인증서 정보는 키를 생성할 당시 사용자가 입력한 정보를 기반으로 고유한 일련번호(Serial Number)가 생성된다. 그리고 이 일련번호는 Java SDK의 'keytools.exe' 도구를 통하여 확인할 수 있다[17].

피해 사례를 분석한 결과, 인증서 파일의 해시는 총 35건의 사례만 중복되어 발견되었다(총 151종류). 이는 사건을 프로파일링하기 위한 정보로 활용하기 어렵다.

그러나 인증서 파일의 일련번호가 중복된 사례는 총 151건으로 약 89%가 재사용되고 있다. Table 6은 인증서 파일에 사용된 상위 10개의 일련 번호 내역이다.

Table 6에 나타나듯이 범죄조직이 악성코드 개발 단계에서 인증서 파일의 일련번호와 퍼미션 파일을 굉장히 높은 비율로 재사용한다는 것이 나타났다. 이를 통하여 위에 열거한 공격도구정보(SMS, 단축 URL, apk 해시, C&C IP, 퍼미션 해시, 인증서 파일 일련번호)를 통하여 기존과 다른 방법의 스미싱

Table 6. Serial number in CERT.RSA

	serial number (in CERT.RSA)	frequency
1	76ad57ed	43
2	936eacbe07f201df	18
3	5cee7bf4	11
4	50fe94fa	10
5	51ebb7fd	8
6	51b47f42	8
7	1a803364	7
8	70114e4d	5
9	5112818e	5
10	512e116b	5

프로파일링 모델 설계를 다음과 같이 제안한다.

4.3 실제 모델 설계 예시

4.3.1 Gephi를 이용한 시각화

Fig 1은 Gephi[18]를 이용한 클러스터링 결과를 나타낸 그림이다. 각 Nodes는 공격도구 정보인 SMS 메시지, apk 파일 해시, 퍼미션 파일의 해시, 인증서 파일의 일련번호이며, 각 Edges는 위의 공격도구 정보를 기반으로 연결하였다.

시각화 결과, 아래와 같이 각 Node들 중 인증서 파일의 일련번호가 가장 많은 빈도로 나타났다. (76ad57ed, 936eacbe07f201df)

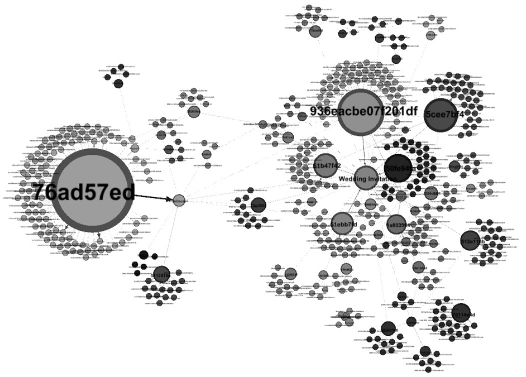

Fig. 1. Visualization of smishing (by Gephi)

4.1.1 공격도구정보를 통한 프로파일링 모델 설계

시각화 결과 인증서 파일의 일련번호가 가장 유의미한 요소로 나타나 이를 가장 큰 비중으로 하여 사건을 군집화(clustering)하였고, 공격도구정보 등을 포함하여 Fig 2와 같이 사건을 군집화하였다.

가장 많은 빈도를 나타낸 인증서 일련번호를 기반으로 10개의 조직으로 군집화하였고, '76ad57ed'에 해당하는 사건이 43건으로 나타났다. 또한, 인증서 일련번호가 다르더라도 퍼미션 파일의 해시가 같은 경우 함께 클러스터링하였다. 그 결과, 일련번호 '5cee7bf4'에 해당하는 사건이 11개에서 16개로 증가하였고, 일련번호 '1a803364'에 해당하는 사건은 기존 7개에서 15개로 증가하였다. 이에 따라, 169개의 사건을 10개의 그룹으로 군집화하여 총 136건을 분류하여, 기존 apk 파일의 해시와 C&C IP로

No.	Serial No.	Hash(AndroidManifest)	SMS Contents	File Name(APK)	HASH(APK)
1	76ad57ed	21c75fa18ebed9cfb58cdbfa2d17bc97	Coupon	SmartBill.apk	10d05330f63ad428ac3ab5d55e9aa1ae
1	76ad57ed	21c75fa18ebed9cfb58cdbfa2d17bc97	Coupon	Smartbill.apk	c23b62750762e3efbed421f5f6e47783
1	76ad57ed	24aff5cb2db1c4efd6de83869f3418f7	Coupon	lotteria.apk	93013e44e56be9a4cf364a6f9d8082be
1	76ad57ed	fc83537e7dd8ad264639bf8cad4f323f	Coupon	pizzahut.apk	a4c60b4eb26172bfcf3b2f06b80fcbb1
1	76ad57ed	fc83537e7dd8ad264639bf8cad4f323f	Coupon	baskin.apk	bd69298a76e8d8c75fe759d641e3992b
2	936eacbe07f201df	1e093f5e961eb789f2e80cdfae947873	Invitation	naver-skt-love.apk	DAC228E629D7C392ECFBC2F3A1BF8C3D
2	936eacbe07f201df	2a832813c2c8709d20e9c84531da1c1a	Invitation	GooglePlayServices.apk	4FE61E218E3120799D72806ABEDF780B
2	936eacbe07f201df	ba1dce17d79715a600801c25718f406b	Invitation	gnppcbank.apk	A7965657F9124E83CFCFF3A6219AE753
2	936eacbe07f201df	c509f5be5b00c7bbfb23781cec2581ae	Android	install.apk	3ef68a5c57f09666fa1d70e7d3eaa131
2	936eacbe07f201df	c509f5be5b00c7bbfb23781cec2581ae	Android	update.apk	58c6ca46280f1562c0d288074cde9e46
3	38203fee	db2a54dec040aa5a35cf73eb9b81f2e5	Payment	siren24app.apk	77d3d2b8f7689c8ba6056e2e85c7ffe1
3	5cee7bf4	038790fd6af66e445bcfb1c5599f9c28	Invitation	sc126114.apk	f23b0d79fbd81569ce704b31fac3fa18
3	5cee7bf4	db2a54dec040aa5a35cf73eb9b81f2e5	Invitation	c12615121.apk	22c18592fa6ed167143aba8a30b62112
3	7c94fc39	db2a54dec040aa5a35cf73eb9b81f2e5	Coupon	CoponFile.apk	3ffe769490f566e43b4337270054c6ab
3	fc61ce7	db2a54dec040aa5a35cf73eb9b81f2e5	Invitation	c160.apk	07EAECE425DE10BB4EEC924649DF4FC5
4	50fe94fa	370fd5ccd708b79ed8e062e7c808f148	Invitation	danal.apk	27D0F10B65D046EB5B236604CAE9AEC8
4	50fe94fa	370fd5ccd708b79ed8e062e7c808f148	Invitation	danal.apk	ba1863d29382d9109db9d060721bc8cd
4	50fe94fa	370fd5ccd708b79ed8e062e7c808f148	Invitation	danal-20130627134803.apk	cd0e468ef38d403f32eb14a0ef16ea53
4	50fe94fa	a483c8ac1d4e8752ffd93c2d03adf592	Invitation	full.apk	0B0A0B45B55D1DD1F4094189F432C741
4	50fe94fa	a483c8ac1d4e8752ffd93c2d03adf592	Invitation	full.apk	AD244D76B9CDB4D2CE30319EBC41ECC4
5	4ffed62b	7b4020d4503289a95a676187ba4b6aef	Invitation	cksoft45.apk	eac190ec79fbbe1e0c6dfd0cff44acba
5	51ebb7fd	7b4020d4503289a95a676187ba4b6aef	Invitation	30.apk	15C10F5747D7C7C6A6F55FBBBD70AD73
5	51ebb7fd	7b4020d4503289a95a676187ba4b6aef	Invitation	wd.apk	3F74790FAF2EC5996378B8F38E1810E2
5	51ebb7fd	7b4020d4503289a95a676187ba4b6aef	Invitation	Wedding.apk	920DD1871DFB5441E8B825C3B373736C
5	51ebb7fd	7b4020d4503289a95a676187ba4b6aef	Invitation	peizhi.apk	F1E23152A7B70C8D402D2677B1B6B59C
6	51828119	7edb97e5eef660cbf83e95246f979c48	Invitation	cheo.apk	cb06f91af37647e92870d81173826fc9
6	51b47f42	7edb97e5eef660cbf83e95246f979c48	Invitation	meena.apk	1a8dce475229d73eeedacd9d43d44513
6	51b47f42	7edb97e5eef660cbf83e95246f979c48	Invitation	enjueng.apk	45E0B9F7EC7A4692C9CD4969B57ACBD5
6	51b47f42	7edb97e5eef660cbf83e95246f979c48	Invitation	yoona.apk	6a233b8fd01b7f5ecfebce1cadb49970
6	51b47f42	7edb97e5eef660cbf83e95246f979c48	Invitation	doen.apk	e3c045da7e74926ae9e2c543d6415805
7	1a803364	1146836b9059ce4e9a572aec570e8b6c		danal.apk	7ba2966d439cb592107bf5e2e1009caf
7	1a803364	f6021c5cda569cce16722d19416f886f	Invitation	danal.apk	f98807fd35fc85622d4a834414a63b9b
7	39e1b5e2	f6021c5cda569cce16722d19416f886f	Invitation	mrk.apk	f057c86cdf69d91fd217995012d59cc2
7	4536af34	f6021c5cda569cce16722d19416f886f	Invitation	danal.apk	15c789b60fa5720457a9a7cf2e1b171d
7	6d296001	f6021c5cda569cce16722d19416f886f	Invitation	danal.apk	799ab23cc70a4c912a797065add77a9d
7	704c92f7	1146836b9059ce4e9a572aec570e8b6c	Invitation	danal.apk	1a2e5db984d5548e5dc4b633c4c763b0
8	70114e4d	6c01c18eda2935af950ce8db3bacf8d4	A Summons	9337y9WdYzkzNGhLMmFyd	2514B5308814B3DA475911ED4BB31171
8	70114e4d	6c01c18eda2935af950ce8db3bacf8d4	Payment	[illegible]	A3BC4A8DB0DAD329AA54E185739A13E6
8	70114e4d	6c01c18eda2935af950ce8db3bacf8d4	A Summons	dH6ikKPE6dkIta3R3OW5TN	fb09a8f783dda28dbc476a3d05dcade4
9	5112818e	0da60537b8308cfa2bc4f84e1bba6670	Coupon	103.apk	4b0034a6c8bc7843e5505a956239483a
9	5112818e	0da60537b8308cfa2bc4f84e1bba6670	Coupon	sm1_210.209.70.213_v1.2.2.a	e73935d8790b4c3681d1e571522ddc52
9	5112818e	0da60537b8308cfa2bc4f84e1bba6670	Coupon	sm1_210.209.70.213_v1.2.2.a	e73935d8790b4c3681d1e571522ddc52
10	512e116b	4edf0b7ae9b6ee413dc8fb08a9d2cbb4	Invitation	m.apk	160E7F2725D1B34495C759845FE1E643
10	512e116b	8d41bffc75738bbf933aa65a95868205	Invitation	luam_c-20130717195154.ap	88EA8E290EC609B0FBDE41AF65131D0F
10	512e116b	d7bba7edca29f988c4e13693dd88a39d	Invitation	google_play_c.apk	e449c4c4264083ed6590736fbc0afd4b

Fig 2. Result of clustering by smishing attack tool information

분류하는 방법보다 큰 효과가 있음을 알 수 있다.

또한 그림과 같이, 군집화를 실시한 결과 SMS 내용 또한 조직별로 거의 유사하게 나타나는 등의 효과를 확인할 수 있었다(무료쿠폰/초대장/출석요구).

4.1.2 Androguard를 이용한 코드유사도 검증

하지만 위의 정보만을 가지고 동일한 조직이 한 범죄임을 결론내리는 것은 쉽지 않다. 그러므로, 코드 유사도 검사를 통하여 좀 더 검증을 강화하는 방법으로 사용한다.

이와 관련하여 2011년 Blackhat 컨퍼런스에서 공개된 Androguard는 안드로이드 응용프로그램 중에서 악성 응용프로그램을 가려내고 이를 분석하는 도구다. 특히 Androguard 중 Androsim을 이용하면, 두 개의 apk 파일의 유사도를 퍼센트(%)로 간단히 나타내준다[19].

인증서 파일의 일련번호와 퍼미션 파일의 해시라는 부분적인 정보를 통하여 대강의 조직을 구별해냈다면, 해당 사례에 사용된 apk 파일의 코드를 서로 비교해봄으로써 그 가능성을 높이는 데 사용한다.

Androsim으로 코드유사도를 비교할 샘플은 Table 7과 같이 선정하였다. 2개의 사례는 겉으로 보기엔 SMS 내용도 다르고, 파일 이름과 파일의 해시 값(MD5), C&C IP가 모두 다르다. 기존의 분류방법에 따르면 이 두 개의 사례는 서로 다른 범죄행위로 간주되었다.

Table 7. Case similarity (two cases)

	1	2
SMS contents	payment	free coupon
File name	siren24app.apk	CoponFile.apk
MD5 (apk)	77d3d2b8f7689c8ba6056e2e85c7ffe1	3ffe769490f566e43b4337270054c6ab
C&C	110.34.237.x	61.176.223.x
Permission	db2a54dec040aa5a35cf73eb9b81f2e5	db2a54dec040aa5a35cf73eb9b81f2e5
Serial number	38203fee	7c94fc39

하지만 새로운 기준에 따라 인증서 일련번호와 퍼미션 파일의 해시로 사례를 군집화하고, 인증서 일련번호가 다르더라도 퍼미션 파일의 해시 값이 동일하므로 같은 그룹의 사례로 가정하였다.

이를 검증하기 위하여 Androsim을 통하여 Fig 3과 같이 코드유사도를 검증하였다. 검증 결과, 두 개의 파일 코드 내 동일한 요소(elements)가 808개, 유사한 요소가 2개로 99.973024%의 유사도가 있는 것으로 나타났다.

즉 본 연구 내용과 같이, 인증서 파일의 일련번호와 퍼미션 파일의 해시 값을 통해 군집화를 하는 작업이 기존 방법에 더해 효과적으로 사용될 수 있는 방법임이 증명되었다. 이렇게 전국에서 발생한 피해사례에 대하여 그룹화를 실시하고(귀납적 프로파일링), 추후 발생하는 사건에서 프로파일링의 각 요소들을 판단하여 기존의 사건에 연결하면(연역적 프로파일링), 효과적인 프로파일링을 할 수 있을 것이다.

```
C:₩Androguard>androsim.py -i CoponFile.apk siren24
app.apk
warning: compressor SNAPPY is not supported (use z
lib default compressor)
Elements:
        IDENTICAL:      808
        SIMILAR:        2
        NEW:            0
        DELETED:        0
        SKIPPED:        0
warning: compressor SNAPPY is not supported (use z
lib default compressor)
        --> methods: 99.973024% of similarities
```

Fig 3. Verification of code similarity by Androsim

V. 결 론

위 모델은 결국 정보분석을 담당하는 중앙 부서에서 관리하여야 하는 요소이다. 프로파일링 기법을 사용하는 것은 결국 '기존의 범죄 사례들을 가지고 특정한 범죄 데이터베이스를 만들어낼 수 있느냐'는 문제와(귀납적 프로파일링) '기존의 범죄 데이터베이스를 가지고 판단했을 때, 새로운 범죄가 기존의 범죄와 유사하다고 볼 수 있느냐'(연역적 프로파일링)는 두 가지의 문제가 있다.

현재 사이버 범죄와 관련된 범죄조직의 데이터베이스가 없기 때문에, 귀납적 프로파일링을 하는 것이 범죄정보분석(intelligence)을 담당하는 부서의 우선순위가 된다. 그리고 이에 대한 검증은, 실제 수사를 담당하는 부서와 프로파일러들의 끊임없는 연구와 노력으로 달성할 수 있을 것이다.

결국, 현장에서 입력한 범죄정보와 악성 애플리케이션 파일을 신속하게 중앙 부서에서 발송 및 관리할 수 있는 시스템을 마련하고, 이를 통하여 신속하게 사건을 묶어 전담팀에게 배당, 범죄자를 추적해야 할 것이다.

본 연구에서는 실제 접수된 169개의 사건정보를 사용하였다. 이를 통하여 전국에 접수된 피해사례를 분석하고 그룹화하는 것이 더 의미 있는 자료가 될 것이다. 이를 통한 검거 사례가 없기에 이를 증명하는 것이 한계점으로 작용하였지만, 추후에는 실제 범인이 검거되었을 때 이 프로파일링 기법이 의미가 있었는지를 판단해 볼 수 있을 것이다.

모바일 악성코드의 분류와 더불어 스미싱 범죄 현상을 이해하고, 범죄조직의 추적을 위한 정보분석능력과 대응능력을 강화하기 위한 스미싱 프로파일링 모델을 제시하였다.

본 연구에서 제안하는 스미싱 프로파일링 기법을 이용한다고 해서 범인을 곧바로 추리할 수 있는 것은 아니다. 단, 사이버범죄는 국경과 관할없이 이루어지고 있으며, 점점 조직화되고 지능화되어 가고 있다. 스미싱 범죄의 피해자를 통해 알 수 있는 정보는 고전적인 수사방식으로는 매우 제한적이다. 그러므로 효율적이고 신속한 수사를 위해서 프로파일링 기법을 통하여 각각의 사례를 유형화 혹은 그룹화해야 한다.

또한 향후 연구에서는 공격도구정보를 통한 프로파일링을 넘어서, 수사가 진행되며 얻을 수 있는 추가적인 정보(공격자정보 및 사건관계인 정보)를 통한 프로파일링 기법을 진행할 예정이다.

References

[1] Korean Nation Police Agency, http://www.police.go.kr/cmm/fms/FileDown.do?atchFileId=FILE_000000000064121&fileSn=1&bbsId=B0000011

[2] Financial Services Commission, http://www.fsc.go.kr/info/ntc_news_view.jsp?menu=7210100&bbsid=BBS0030&no=29505, Dec. 2013.

[3] InfoWorld, http://www.infoworld.com/article/2659058/security/mcafee-warns-of-smishing--attacks.html

[4] Korean Nation Police Agency, http://www.police.go.kr/portal/main/contents.do?menuNo=200287

[5] Jae-sung Yun, et al. "Andro-profiler: Anti-malware system based on behavior profiling of mobile malware," *Journal of The Korea Institute of Information Security & Cryptology*, 24(1), pp. 145-154, Feb. 2014.

[6] Changwook Park, et al. "Research on the Classification Model of Similarity Malware using Fuzzy Hash," *Journal of The Korea Institute of Information Security & Cryptology*, 22(6), pp. 1325-1336, Dec. 2012.

[7] Jae-woo Park, et al. "An Automatic Malware Classification System using String List and APIs," *Journal of Security Engineering*, 8(5), pp. 611-626, Oct. 2011.

[8] Dong-Jie Wu, et al. "Droidmat: Android malware detection through manifest and API calls tracing," *Information Security (Asia JCIS), 2012 Seventh Asia Joint Conference on*. IEEE, 2012.

[9] Choon Kyong Joo and Ji won Yoon, "Discrimination of SPAM and prevention of smishing by sending personally identified SMS(For financial sector)," *Journal of The Korea Institute of Information Security & Cryptology*, 24(4), pp. 645-653, Aug. 2014.

[10] SangKeun Jang, "A strategy for mobile malicious code and a method for diagnosis of moblie malicious code by case anlaysis," *KIISC Review*, 23(2), pp. 14-20, Apr. 2013.

[11] Douglas, John E., et al. "Criminal profiling from crime scene analysis," Behavioral Sciences & the Law, vol. 4, no. 4, pp. 401-421, 1986.

[12] Chaeho Lim, et al. "Profiling of Cyber-crime by Psychological View," *Journal of The Korea Institute of Information Security & Cryptology*, 19(4), pp. 115-124, Aug. 2009.

[13] Cheol-Woo Jeong, et al. "A design for Profiling-system of Cyberterrorism," Korean National Police Agency, Jan. 2013.

[14] Hongsuk Yang, "A model design for Profiling of DDoS Crime," Ph.D. Thesis, Korea University, Feb. 2012.

[15] Neumann, Alexander, Johannes Barnickel, and Ulrike Meyer. "Security and privacy implications of url shortening services," Proceedings of the Workshop on Web 2.0 Security and Privacy, 2010.

[16] Android Developer, "app-signing", http://developer.android.com/tools/publishing/app-signing.html

[17] Java SE Documentation, "keytool", http://docs.oracle.com/javase/6/docs/technotes/tools/windows/keytool.html

[18] Gephi v0.8.2, http://gephi.github.io/

[19] Jun-Hyung Kim and Eul-Gyu Im, "Androguard: Similarity Analysis for Android Application Binaries", *Korea Computer Congress 2014*, pp. 101-103, Jun, 2014.

〈저자소개〉

정 영 호 (Youngho Jeong) 정회원
2009년 3월: 경찰대학 행정학과 학사 졸업
2013년 7월~2015년 2월: 고려대학교 정보보호대학원 석사
2014년 2월~현재: 경찰대학 국제 사이버범죄 연구센터 연구원
〈관심분야〉 디지털 포렌식, 사이버범죄, 정보분석

이 국 헌(Kukheon Lee) 학생회원
2012년 2월: 배재대학교 컴퓨터공학 공학사
2012년 3월~2014년 8월: 고려대학교 정보보호대학원 석사
2014년 9월~현재: 고려대학교 정보보호대학원 박사과정
〈관심분야〉 디지털 포렌식, 데이터베이스 포렌식

이 상 진 (Sangjin Lee) 종신회원
1989년 2월~1999년 2월: 한국전자통신연구원 선임 연구원
1999년 2월~2001년 8월: 고려대학교 자연과학대학 조교수
2001년 9월~현재: 고려대학교 정보보호대학원 교수
〈관심분야〉 대칭키 암호, 정보은닉 이론, 컴퓨터 포렌식